

## **Developing a North-American Strategy to Preserve & Manage Print Collections of Monographs**

This document serves as background for workshop participants. The vision it sketches is not a blueprint for a completed program but rather the scaffolding on which the workshop will build a plan that responds to needs, trends, and alternatives; lays out components of a system; and defines the work that needs to be done.

### ***Statement of Problem and Workshop Background***

As libraries divest of older, unused materials, repurpose local library space, and address decreasing financial resources to manage collections and facilities, it is imperative that they develop a shared approach to cooperative management of legacy monograph collections. Libraries are already working toward collaborative management of journals, newspapers, legal materials, and government documents, but concerted efforts focused on print monographs have not emerged, due in part to the complex challenges presented by monograph collections.

With support from the federal Institute of Museum and Library Services (IMLS), LYRASIS and its planning partners the Center for Research Libraries (CRL), the Committee on Institutional Cooperation (CIC), and the California Digital Library (CDL) are sponsoring a national discussion to explore a framework for collaborative retention of print monograph collections.

The key feature of the six-month planning initiative is a multi-day 'think tank' that will convene library leaders from across the country to define characteristics of a collaborative monograph archiving model and to identify issues that require testing or research to confirm a framework for future action. In focusing intense attention by participants on the possibilities for collaborative retention and preservation of print collections of monographs, the workshop builds on the approach successfully used in the 2003 "Preserving America's Print Resources" (PAPR) workshop sponsored by CRL and funded by IMLS, which sparked discussion and action on a number of formats, particularly journals.

### ***Workshop Outcomes***

The workshop will produce two outcomes:

- 1) Recommendations for a nation-wide strategy for shared print monograph retention and preservation.
- 2) A list of research or test projects needed to confirm the strategy; these projects will form the basis of one or more future grant proposals.

### ***Seeding Discussion Through a Vision: Outlining an Ideal Framework***

Development of the framework requires a common understanding of the environment in which libraries work, the experiences that libraries have had in addressing preservation and retention of collections, the communication and record management opportunities afforded by technology, and the complexities of developing a national strategy to preserve rich monographic collections. That common understanding most likely includes the following assumptions:

*In the next 5-10 years, mass digitization projects will mature and their attendant legal issue be addressed; the majority of staff and faculty will employ congenial software and devices to read and use most texts in their digital version (at least in the first instance of reading); libraries and their home institutions will continue to redevelop library spaces for purposes other than housing of printed materials; few libraries (and even fewer parent institutions and funding sources) will want to create more storage space off campus to house their own collection; and cultural changes resulting from online bookselling, user-initiated borrowing through regional/national catalogs of library holdings, discovery of books on the Web, and POD technologies will persuade most users that most tangible versions of library materials can be delivered in a reasonable period of time rather than held in open stacks on campus.*

*These changes will occur faster for some fields and user groups than others. As trusted stewards of the record of scholarship, libraries will continue to work collaboratively to capitalize on their long history of resource-sharing, ensuring that printed works are preserved in adequate numbers as insurance against the need for re-digitization and to afford ongoing access to texts in the form in which they were published.*

*In developing this stewardship program libraries will build on their experiences with collaborative approaches to preservation of and access to print journal backfiles and other types of material, approaches which demonstrate the importance and power of building a strong structural framework that includes well-defined goals, organizational responsibilities and relationships, and policy and governance arrangements. Libraries will also consider the lessons learned from emerging collaborative models focused on monographs.*

The ideal framework must develop over the next few years a structure for legacy print monographs that accomplishes the following:

- helps libraries collectively preserve the record of scholarship published in monographs
- creates a systematic, coordinated, sustainable, and strategic approach that replaces local, ad hoc, and independent approaches to de-accessioning or storage of monographs
- develops a process that identifies stakeholders, analyzes opportunities and vulnerabilities, and builds on patterns of relationships, recognizing the need for libraries to repurpose space and achieve savings in housing costs

by reducing unnecessary duplication while preserving an adequate number of copies

- uses an information system that discloses retention decisions and responsibilities, facilitates large-scale holdings comparisons of print and digitized monographs, and automatically generates reports for libraries of items they should retain or may consider for removal
- provides rapid access, when needed, to intentionally retained copies
- supports discoverability of print copies and digital surrogates in the same discovery layer
- provides avenues for a broad spectrum of libraries to financially support and sustain the retention commitments and access services

### *Developing the Vision, describing the Strategy*

The following five topics will constitute the primary agenda of the workshop. Participants will be divided into working groups for these topics based on expressed interest (participants may be assigned to more than one group). Plenary and breakout sessions will allow participants to interact on all topics.

The five topics are as follows:

#### *A. Enumerate the archiving/preservation issues that are specific to monographs.*

What is the problem to be solved and what are the expected benefits? How is preservation of monographs different from preservation of serials? It will be important to identify any special considerations related to monographs so that collaborative archiving, preservation, and access programs may be designed to address them.

For example:

- The sheer numbers: A 2009 study of OCLC's WorldCat determined that 84.8 million of its 135.3 million records represented printed monographs, 63% of the database;
- High uniqueness rates: Of the 26 million distinct print monograph titles identified in a 2005 study of OCLC's WorldCat, 36% are uniquely held and less than 10% are held by 50 or more institutions<sup>2</sup>
- Differences in imprints and editions that vary with the date/place of publication, posing questions about which variants to retain;
- Differences in condition, binding, marginalia, and other physical characteristics that make it more likely for any given copy to be damaged or otherwise marked, requiring an archiving arrangement to ensure that copies retained are sound and complete;
- Differences in the quality and completeness of metadata describing monographs, including potentially lack of metadata and/or holdings information for many monograph titles;

- Broad engagement of organizations in digitization of monograph titles, with wide variance in quality, lack of clarity regarding preservation strategies for the digital surrogate, and limited access to metadata;
- Challenges with delivery to users: Digital delivery of journal articles from offsite storage is common practice, but it is more costly and not always feasible for monographs; at the same time, physical delivery of monographs may not be cost-effective in a national model.
- De-selection of monographs is dauntingly labor-intensive if undertaken title-by-title and, unlike de-selection of serial publications, can be inefficient in terms of the space gained per decision made.

It will be useful to suggest ways to break the monolithic term "monograph" into categories or types of publications in order to address the needs of each according to a risk assessment framework and to prioritize these types or categories of publications for action. For the purposes of this work, a monograph may be defined as a non serial publication complete in one volume or a finite number of volumes. As a pragmatic definition, a monograph might more easily be identified by the OCLC catalog record/MARC fixed field BLvl code "a." Are these definitions appropriate?

*B. Describe the relationship between and issues related to digital surrogates and print archive copies.*

Suggest a scenario (or scenarios) in which the availability of digital surrogates will become the library collection of first resort. What conditions need to be in place in order to allow digitized copies to be the primary means for accessing and using texts? What do we know now about how the existence of digitized text affects use of print and what more do we need to know? Recognizing the importance of linking digital surrogates to their archived monographs, how will we design a workable model to associate these variant formats, addressing any legal issues as well as bibliographic details that may arise. A strategic starting point might be to identify titles in the HATHI Trust compared with printed monographic titles that are unique or scarce.

*C. Outline the requirements for a bibliographic information/disclosure system (or alternative systems) that would enable large-scale collaboration on monographs among libraries.*

A nation-wide framework must describe the ideal information system to support selection and retention of archived monographs.

A sample scenario might be described thusly:

*For a library to be a good citizen in preserving the scholarly record as it divests of holdings it no longer wants, it would need in its catalog records a note that identifies unique or scarcely held titles at the regional or national level. This note would be used to produce reports as libraries make divestment decisions. When a library divests of*

*batches of titles, its holdings symbols would be taken off OCLC and regional catalogs in batches. When a library keeps unique or scarcely held titles, that information is batch loaded to OCLC or regional catalogs. A library could automatically send unique or scarcely held books it is divesting of to a regional repository that does not have them among its holdings; appropriate information about such batch transfers would go into local, regional, and national catalogs in batch at the same time.*

As a starting point, how well do the following statements describe the information system needed to support collaborative monographic archiving? What other components are needed and how would they work?

- be as automated as possible so that a very high number of individual disposition decisions can be made by machine;
- clearly and broadly communicate holdings info (in local as well as national catalogs) about the copies in a national print archiving program;
- employ a system that allows batch comparison of holdings (what has ReCAP learned about mass comparison?);
- enable title/copy-level decisions in large numbers automatically, holdings changes in national catalog automatically, and automatically enable acceptance of scantily-held discards by a depository;
- use standards for verification of condition like used-book dealers have; condition goes into note about stored/archived copies;
- enable selection of titles for digitization and feed scarcely held titles into the digitization stream.

Significant work is underway among print journal archiving projects (such as WEST and CRL) to define disclosure methods, metadata standards, and collection analysis functions as described above. Can monograph archiving projects build on or participate in this work? Are there requirements specific to monographs which must be accommodated?

*D. Outline the characteristics of service models and business models to sustain retention commitments and provide new modes of discovery and access to retained copies.*

We know that the future environment for research collections is likely to include dynamic distribution permissions for monographs that acknowledge the evolving interpretations and landscape of copyright laws. A section of the monograph collections today may be in copyright and the digital surrogates not distributable, it may be orphaned and possibly distributable electronically, or it may be in public domain and distributable. In all cases, the print copies for those sections of the collections are also distributable. The future environment is likely to include user preferences for print and electronic depending upon the distribution permissions at a given point in time (and other factors). Users may prefer to discover both formats in the same discovery layer in order to make choices about which format to access and their preferred turnaround/access time. Users also may prefer rapid delivery to the physical and virtual space in which s/he conducts research.

In this likely environment, the workshop planners start with several assumptions about an eventual print monographic collections framework. They are, of course, open to

discussion and are based on recent projects with other formats and on the history of collaboration on monograph collections. The framework will exhibit some or all of the following characteristics:

- rely on and take advantage of existing relationships and programs established by local and regional consortia or other cooperatives and organizations;
- rely initially on copies that are already in storage;
- avoid creating an entity to oversee a program and instead will rely on an organization that already exists, one that will agree to serve coordination and communication functions for regional projects;
- begin as a "what not to withdraw" program that identifies and protects the unique and scarce;
- not establish an agreement regarding the minimum number of copies that should be archived in all cases but will rely on guidelines for a more ad hoc approach to copy retention and disposal, specifying only the general conditions and protocols for local and regional determination of which library keeps which copies;
- rely on a voluntary declaration by libraries of their willingness to archive copies; such libraries will probably be those that have historically regarded their mission's being to preserve the print record;
- need flexible preservation policies and storage requirements that include desiderata for treatment of damaged volumes and state minimum environmental controls and monitoring and disaster plans;
- have back-up and sunset strategies in case a participant needs to withdraw from the program;
- be flexible with respect to the nature and level of access to archived monographs (how "light" or "dark" should stored collections be);
- suggest a trajectory for moving a monograph from storage to "true" archival status.

With these assumptions about the framework in mind, a likely future environment for monographs may include, then, fewer physical copies retained by some research libraries on behalf of a network of libraries. Their doing so will require new business models to sustain those retention commitments, new engagement in collection development with those supporting libraries, and new modes of access services to provide rapid delivery to the physical space in which the researcher works (and/or the nearest research library.)

What are the desiderata of a service model and business model for preserving print monographs and providing access to them in a multiformat, dynamic distribution rights environment?

Requirements to be considered include:

- Incorporate the mechanics of the retention/disclosure system
- Include a retention and collection development decision framework specific to print monographs at the network level
- Provide discovery of retained print monographs and digital surrogates (when they exist) in the same discovery layer at the network level

- Show the user how to move from the network level discovery layer to access to materials
- Access policies for circulation, lending, and duplication within the collective archive
- Include an access service that allows the user to choose which format to receive based on preferred delivery time (download, standard, expedited), format (print or electronic) and delivery location (hard drive, handheld device, office, home or nearest research library to which s/he belongs)
- Recover costs for retention commitments
- Recover costs for access services to the retained copies
- Include a sustainable model

*E. Design a process for assessment that includes goals and milestones and evaluates the framework on an on-going basis.*

Assessment markers can include a report on how many of the estimated 26 million distinct monograph titles are already in storage; how many unique titles have been digitized and stored; how often testing of accuracy of bibliographic records and items should occur to sustain a trusted system; and reviewing procedural approaches to ensure that the system invites participation and ease of use by libraries and scholars. The business model requires consistent review for sustainability, sufficient economic leveraging for networks and individual libraries, and the on-going dynamics of print monographs and digitization projects.

<sup>1</sup> Brian Lavoie and Lorcan Dempsey, "Beyond 1923: Characteristics of Potentially In-copyright Print Books in Library Collections," *D-Lib Magazine* 15:11/12 (November/December 2009). Available at [www.dlib.org/dlib/november09/lavoie/11lavoie.html](http://www.dlib.org/dlib/november09/lavoie/11lavoie.html) (retrieved 20 November 2009).

<sup>2</sup> Brian F. Lavoie and Roger Schonfeld, "A Brief Tour of the System-wide Print Book Collection," *Journal of Electronic Publishing* 9:2 (Summer 2006). Available at <http://quod.lib.umich.edu/cgi/t/text/textidx?c=jep;cc=jep;q1=3336451.0009.2%2A;rgn=main;view=text;idno=3336451.0009.208> (retrieved 16 December 2009).